

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2003-032680

(43)Date of publication of application : 31.01.2003

---

(51)Int.Cl.

H04N 7/30

H04N 1/41

---

(21)Application number : 2001-218665 (71)Applicant : CANON INC

(22)Date of filing : 18.07.2001 (72)Inventor : OOSA KINYA

---

(54) IMAGE CODED, IMAGE CODING METHOD, PROGRAM CODE OF THE SAME IMAGE CODER, AND MEMORY MEDIUM OF THE SAME IMAGE CODER

(57)Abstract:

PROBLEM TO BE SOLVED: To reduce the difference between the image quality of each tile and the one of input image of a image coder, even when decoding the coded data of each tile.

SOLUTION: A code-quantity control portion 104 of the image code so decides every tile divided by a tile dividing portion 101 the basic code-quantity to be assigned to each tile as to control every tile the increase/decrease of the basic code-quantity to be assigned to each tile, by using the information of the flatness calculated every tile by a tile-flatness calculating portion 17.

---

## LEGAL STATUS

[Date of request for examination] 18.11.2003

[Date of sending the examiner's decision of rejection] 04.08.2006

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right]

---

## CLAIMS

---

[Claim(s)]

[Claim 1] A division means to be image coding equipment which encodes an image according to JPEG2000 (ISO/IEC 15444-1) specification, and to divide an image into a tile, A conversion means to perform transform processing to each tile and to generate a transform coefficient, A coding means to perform coding processing to said transform coefficient, and an amount count means of basic signs to calculate the amount of basic signs assigned for every tile, An index count means to calculate the index which shows the display flatness of an attention tile based on the data of an attention tile, According to the index which shows the display flatness of the attention tile calculated by said index count means in the amount of basic signs assigned to the attention tile calculated by said amount count means of basic signs, so that display flatness is large and an allotment may increase Image coding equipment characterized by having an amount count means of correction signs to make an allotment fluctuate, and the control means which controls said coding means based on the amount of allocation codes of the attention tile calculated by said amount count means of correction signs.

[Claim 2] Said index count means is image coding equipment according to claim 1 characterized by using the inverse number of distribution of the data of said attention tile as an index of said attention tile.

[Claim 3] Said index count means is image coding equipment according to claim 1 characterized by the thing of difference with the average of each data of said attention tile, and each data concerned for which it asks for the sum absolutely and the inverse number is used as an index of said attention tile.

[Claim 4] Said index count means is image coding equipment according to claim 1 characterized by asking for total of the absolute value of the difference of the data of two \*\*\*\*\*, and using the inverse number as an index of said attention tile when each data of said attention tile is scanned in order.

[Claim 5] The data of said attention tile are image coding equipment given in claim 1 characterized by being the value of the pixel contained in said attention tile thru/or any 1 term of 4.

[Claim 6] The data of said attention tile are image coding equipment given in claim 1

characterized by being the transform coefficient which constitutes LL subband of said attention tile thru/or any 1 term of 4.

[Claim 7] Said amount count means of basic signs is image coding equipment according to claim 1 characterized by calculating the amount of signs which is to the bases assigned to said attention tile using a rate / distortion optimization.

[Claim 8] Said amount count means of correction signs is image coding equipment according to claim 1 characterized by asking by calculating  $B = B_{base} \times (1 + f(F))$  for the amount B of signs assigned [ as opposed to / when it considers as the function which increases function  $f(x)$  for the fundamental amount of signs which assigns the index of said attention tile to F and said attention tile in monotone to the value of x to  $B_{base}$  / said attention tile ].

[Claim 9] Said amount means of correction signs is image coding equipment according to claim 8 characterized by saturating  $f(x)$  in a upper limit to the big value of x, and saturating  $f(x)$  in a lower limit to the small value of x.

[Claim 10] The division process which is the image coding approach which encodes an image according to JPEG2000 (ISO/IEC 15444-1) specification, and divides an image into a tile, The conversion process which performs transform processing to each tile and generates a transform coefficient, The coding process which performs coding processing to said transform coefficient, and the amount count process of basic signs which calculates the amount of basic signs assigned for every tile, The index count process which calculates the index which shows the display flatness of an attention tile based on the data of an attention tile, According to the index which shows the display flatness of the attention tile calculated at said index count process in the amount of basic signs assigned to the attention tile calculated at said amount count process of basic signs, so that display flatness is large and an allotment may increase The image coding approach characterized by having the amount count process of correction signs of making an allotment fluctuating, and the control process which controls said coding process based on the amount of allocation codes of the attention tile calculated at said amount count process of correction signs.

[Claim 11] The image coding approach according to claim 10 characterized by using the inverse number of distribution of the data of said attention tile as an index of said attention tile at said index count process.

[Claim 12] The image coding approach according to claim 10 characterized by the thing of difference with the average of each data of said attention tile, and each data concerned for which it asks for the sum absolutely and the inverse number is used as an index of said attention tile at said index count process.

[Claim 13] The image coding approach according to claim 10 characterized by asking for total of the absolute value of the difference of the data of two \*\*\*\*\*, and using the inverse number as an index of said attention tile at said index count process when each data of said attention tile is scanned in order.

[Claim 14] The data of said attention tile are the image coding approach given in claim

10 characterized by being the value of the pixel contained in said attention tile thru/or any 1 term of 13.

[Claim 15] The data of said attention tile are the image coding approach given in claim 10 characterized by being the transform coefficient which constitutes LL subband of said attention tile thru/or any 1 term of 13.

[Claim 16] The image coding approach according to claim 10 characterized by calculating the amount of signs which is to the bases assigned to said attention tile at said amount count process of signs using a rate / distortion optimization.

[Claim 17] The image coding approach according to claim 10 characterized by to calculate the amount B of signs assigned to said attention tile by calculating  $B = B_{base} \times (1 + f(F))$  at said amount count process of correction signs when it considers as the function which increases function  $f(x)$  for the fundamental amount of signs which assigns the index of said attention tile to F and said attention tile in monotone to the value of x to Bbase.

[Claim 18] Said amount means of correction signs is the image coding approach according to claim 17 characterized by saturating  $f(x)$  in a upper limit to the big value of x, and saturating  $f(x)$  in a lower limit to the small value of x.

[Claim 19] The program code which performs the image coding approach of a publication in claim 10 thru/or any 1 term of 18.

[Claim 20] A program code according to claim 19 is stored, and it is the storage which a computer can read.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention] This invention relates to a program code and a storage at the image coding equipment which encodes an image and the image coding approach, and a list.

[0002]

[Description of the Prior Art] Like a JPEG2000 image coding method (ISO / IEC FDIS 15444-1), as compared with the JPEG base-line coding method using DCT of the block base etc., even if the coding method which carries out Wavelet conversion of the image and is encoded raises condensation, it does not produce a block strain but has an image quality-merit. Although it is effective to perform Wavelet conversion to the whole input image in order to employ such an image quality-merit in the maximum efficiently In actually mounting coding equipment, with the very large size of an input image Or how to be unable to Wavelet-change the whole input image at once, and to

be unable to encode, but to divide an input image into two or more fields (tile), and to encode each tile according to the reason of \*\* the buffer capacity of coding equipment is restricted, respectively may be taken.

[0003] The example of processing of such tile division is shown in drawing 2. In this drawing, the input image 21 is divided into a tile 22 or a tile 23 by the tile division section 24 in a dotted-line part, and every one divided tiles 22 and 23 are encoded with an encoder. Since size is small compared with the input image 21, the divided tiles 22 and 23 limit the buffer capacity which it has with a next encoder, and enable mounting by low cost more easily.

[0004] The functional configuration of the image coding equipment which was divided into drawing 3 by carrying out tile division of the input image and which encodes for every tile is shown.

[0005] 301 is the tile division section, divides an input image into the field (tile) which has predetermined size, and outputs every one divided tile to the conversion coding section 302. 302 is the conversion coding section, it performs transform processing to the inputted tile, generates a transform coefficient, and outputs it to quantization and the entropy-code-modulation section 303. This transform processing has frequency conversion and has discrete wavelet transform, DCT, etc. as frequency conversion. 303 is quantization and the entropy-code-modulation section 303, quantizes to a transform coefficient first and acquires the quantization value of a transform coefficient. And entropy code modulation is performed to this quantization value, and it outputs as coded data. In addition, a header is attached to this coded data and the parameter used for above-mentioned coding processing, for example, a quantization step, the size of a tile, etc. are contained in this header. 304 is the amount control section of signs, it quantizes in quantization and the entropy-code-modulation section 303, and is supervising the amount of signs generated in case entropy code modulation is performed, and controls quantization and the entropy-code-modulation section 303 according to the amount of signs to generate (coding control). For example, it is made for the size of the coded data to generate to turn into size specified beforehand.

[0006] As the technique of this amount control of signs, it is ISO/IEC. In the specified amount of signs, the technique of the rate / distortion optimization which minimizes the error of the encoded image and an input image is used, and the amount of signs to generate is controlled by the knot of "Rate control" of FDIS 15444-1 so that image quality with the most sufficient error in the semantics of min is acquired.

[0007]

[Problem(s) to be Solved by the Invention] Performing the amount control of signs using the technique of a rate / distortion optimization becomes the control from which the best image quality is acquired in the semantics of min in an error. However, it is not necessarily in agreement that an error is min and that image quality is good for a vision target with human being's visual property. In encoding by performing tile

division especially, it produces the following problems only by minimizing an error. It explains using the input image 21 which showed this problem to drawing 2.

[0008] The tile 22 which includes many flat image fields in the input image 21 first, and the tile 23 including many image fields which are not flat are compared. A flat image field shows the field where each pixel value change contains a comparatively small pixel here. On the other hand, the image field which is not flat shows the field containing a pixel with each comparatively large pixel value change.

[0009] ISO/IEC If the amount control of signs described by FDIS 15444-1 is performed, the amount control of signs will be performed so that the error over the whole image may be minimized, and allocation of the amount of signs optimal in respect of an error also for a tile 22 and a tile 23 will be performed. However, if it becomes the image quality in which the flat tile 22 faded compared with the tile 23 to the thing with the comparatively good image quality of the tile 23 which is not flat when the image after decode was looked at by the eye actually and sees as the whole image, it will become a flat part and an image with the sense of incongruity which has a difference in the image quality of the part which is not flat. This is the effect of the masking effect in human being's vision property, and in the image field which is not flat, in order that masking may not work in a flat field to being gross errors or a coding error being masked by fluctuation of the image information itself, and not being somewhat conspicuous, it is because it is conspicuous also with a small error.

[0010] Even if it accomplishes this invention in view of the above problem and decodes the coded data of each tile, it aims at mitigating the difference of the image quality for every tile.

[0011]

[Means for Solving the Problem] In order to attain the object of this invention, the image coding equipment of this invention is equipped with the following configurations.

[0012] Namely, a division means to be image coding equipment which encodes an image according to JPEG2000 (ISO/IEC 15444-1) specification, and to divide an image into a tile, A conversion means to perform transform processing to each tile and to generate a transform coefficient, A coding means to perform coding processing to said transform coefficient, and an amount count means of basic signs to calculate the amount of basic signs assigned for every tile, An index count means to calculate the index which shows the display flatness of an attention tile based on the data of an attention tile, According to the index which shows the display flatness of the attention tile calculated by said index count means in the amount of basic signs assigned to the attention tile calculated by said amount count means of basic signs, so that display flatness is large and an allotment may increase It has an amount count means of correction signs to make an allotment fluctuate, and the control means which controls said coding means based on the amount of allocation codes of the attention tile calculated by said amount count means of correction signs.

[0013]

[Embodiment of the Invention] With reference to an accompanying drawing, this invention is explained to a detail according to a suitable operation gestalt below.

[0014] The functional configuration of the image coding equipment in this operation gestalt is shown in [operation gestalt of \*\* 1st] drawing 1 .

[0015] 101 is the tile division section, an input image is divided into the field (tile) of predetermined size, and each tile is outputted to the conversion coding section 102. [0016] 102 is the conversion coding section, performs transform processing to each tile, and generates a transform coefficient. With this operation gestalt, frequency conversion is used as this transform processing. In addition, these any are sufficient although DCT and discrete wavelet transform are contained in this frequency conversion.

[0017] 103 is quantization and the entropy-code-modulation section, quantizes using a predetermined quantization step to the transform coefficient by the conversion coding section 102, and performs entropy code modulation to the quantized transform coefficient.

[0018] 104 is the amount control section of signs, supervises the amount of signs of the quantization performed in quantization and the entropy-code-modulation section, and the coded data generated by the entropy-code-modulation section, and controls quantization and the entropy-code-modulation section 104 based on this amount of signs. It mentions later for details.

[0019] 105 is the tile display flatness count section, and outputs the display flatness which calculated and calculated the display flatness for every tile to the amount control section 104 of signs. It mentions later for details.

[0020] Moreover, the basic configuration of the image coding equipment in this operation gestalt is shown in drawing 5 .

[0021] 501 is CPU, and it performs image coding processing of this operation gestalt in which it has the functional configuration shown in drawing 1 while it controls the whole image coding equipment using the program and data which were stored in RAM502 or ROM503.

[0022] 502 is RAM, and it is equipped with the work area used in case CPU501 performs various kinds of processings while it is equipped with the area which stores temporarily the program loaded from external storage 504 or the storage drive 509, data, etc.

[0023] 503 is ROM and stores a program, data, etc. which control the whole image coding equipment.

[0024] 504 is external storage, such as a hard disk, and can save a program, data, etc. which were loaded from the storage drive 509. Moreover, when the size of a work area exceeds allowance of RAM502, the area to have exceeded can be offered as a file.

[0025] 505 and 506 are a keyboard and a mouse and can input various kinds of directions to image coding equipment.

[0026] 507 is a display, is constituted by CRT, the liquid crystal screen, etc. and can display various kinds of image information, text, etc.

[0027] It is a picture input device, and 508 is constituted by the scanner, the digital camera, etc., it can use an image as digital data and can input it into image coding equipment. In addition, a picture input device includes the interface for connecting with image coding equipment.

[0028] 509 is a storage drive, can read a program, data, etc. in storages, such as CD-ROM, DVD-ROM, and a floppy (trademark) disk, and can output them to RAM502 or external storage 504.

[0029] 510 is a bus which connects above-mentioned each part.

[0030] The processing in the amount control section 104 of signs of the image coding equipment of this operation gestalt equipped with the above configuration and the tile display flatness count section 105 is hereafter explained to a detail.

[0031] The amount control section 104 of signs determines the fundamental amount of signs assigned for every tile by technique, such as a rate / distortion optimization indicated by ISO/IEC FDIS 15444-1. And control which fluctuates the fundamental amount of signs to assign for every tile is performed using the information on the display flatness for every tile calculated by the tile display flatness count section 17.

[0032] If it is that the display flatness of Tile  $i$  is given by  $F_i$  (a tile is so flat that  $F_i$  is large), the amount  $B_i$  of signs actually assigned to Tile  $i$  can be obtained by making fundamental amount quota  $B_{base-i}$  of signs for every tile fluctuate according to  $F_i$  like a degree type.

[0033]

$$B_i = B_{base-i} \times (1 + f(F_i))$$

Function  $f(x)$  is a function which increases in monotone to the value of  $x$ , for example, is  $f(x) = a(x-b)$  here. However,  $a$  and  $b$  are constants and can consider  $a > 0$ . In the case of the function of this example, the value of a constant  $a$  determines the degree of the change in the sign allotment  $B_i$  to the variation of display flatness  $F_i$ , and when the value of display flatness  $F_i$  is larger than the value of a constant  $b$ ,  $f(F_i) > 0$ , i.e.,  $B_i$  is made to increase, and when the value of  $F_i$  is smaller than the value of  $b$ ,  $f(F_i) < 0$ , i.e.,  $B_i$  is decreased. In addition, it is determined that the value of constants  $a$  and  $b$  becomes the amount of signs which a visual improvement effect is acquired and the amount of signs as the whole image means.

[0034] In order that the sign allotment  $B_i$  may avoid changing to an extreme value to the very big value of  $x$ , and the very small value of  $x$  although not restricted to the function of this example if it is the function which increases in monotone, as for function  $f(x)$ , it is desirable to saturate the value of function  $f(x)$  with an upper limit or a minimum to such  $x$ .

[0035] Although the inverse number of distribution of each pixel value in Tile  $i$  is used as display flatness  $F_i$  calculated in the tile display flatness count section 105, it is not limited to this, and may ask for the absolute value sum of the difference of each pixel



value in a tile, and the average of a pixel value, and the inverse number may be used. moreover -- or when each pixel value in a tile is scanned in order, it can ask for total of the absolute value of the difference of the 2 pixel value of \*\*\*\*\*, and the inverse number can also be used as display flatness.

[0036] When the average of  $P_{ij}$  ( $0 \leq j < N - 1$  and  $N$  are the number of pixels in Tile  $i$ ) and  $P_{ij}$  was made into  $P'$  for the pixel value in Tile  $i$  and the inverse number of distribution of each pixel value in Tile  $i$  is used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

[0037]  $F_i = N / \{ (P_i0 - P')^2 + (P_i1 - P')^2 + \dots + (P_i(N-1) - P')^2 \}$

On the other hand, when it asks for the absolute value sum of the difference of each pixel value in a tile, and the average of a pixel value and the inverse number is used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

[0038]  $F_i = N / \{ |P_i0 - P'| + |P_i1 - P'| + \dots + |P_i(N-1) - P'| \}$

Or when each pixel value in a tile is scanned in order, it asks for total of the absolute value of the difference of the pixel value of two \*\*\*\*\*, and the inverse number is used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

[0039]  $F_i = N / \{ |P_i1 - P_i0| + |P_i2 - P_i1| + \dots + |P_i(N-1) - P_i(N-2)| \}$

And the amount control section 104 of signs calculates the amount  $B_i$  of signs actually assigned for every tile based on an above-mentioned formula using the display flatness  $F_i$  for which it asked as the tile display flatness count section 105 was a \*\*\*\*.

[0040] If coding processing is performed to all tiles by the above processing, coded data including the coding result of each tile will be generated. Moreover, the header containing the parameter used for the quantization step and entropy code modulation for every tile is attached to this coded data.

[0041] By the above explanation, since the image coding equipment and the image coding approach in this operation gestalt perform coding in consideration of the display flatness for every tile, even if they decode these tiles, they can mitigate the difference of the image quality for every tile.

[0042] The case where wavelet transform is used as frequency conversion processing in the (operation gestalt of \*\* 2nd) conversion coding section 101 is explained. Each tile will be divided into a subband as shown in drawing 4 if the conversion coding section 101 performs wavelet transform for every tile.

[0043] If display flatness is calculated to LL subband even if it does not calculate display flatness to the whole tile since the multiplier of LL subband is equivalent to the cutback image of a tile in this drawing, the display flatness of a tile is computable by small computational complexity. When the average of  $LL_{ik}$  ( $0 \leq k < M - 1$  and  $M$  are the number of the transform coefficients in LL subband) and  $LL_{ik}$  was made into  $LL'$  for the multiplier value of LL subband of Tile  $i$  and the inverse number of distribution

of each transform coefficient in Tile i is specifically used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

$$[0044] F_i = M / \{ (LLi0 - LL')^2 + (LLi1 - LL')^2 + (LLi(M-1) - LL')^2 \}$$

On the other hand, when the absolute value of the difference of the average of each transform coefficient in the subband LL and a transform coefficient is calculated and the inverse number is used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

$$[0045] F_i = M / \{ |LLi0 - LL'| + |LLi1 - LL'| \text{ and } + |LLi(N-1) - LL'| \}$$

Or when each transform coefficient value in the subband LL is scanned in order, it asks for total of the absolute value of the difference of the transform coefficient value of two \*\*\*\*\* and the inverse number is used as display flatness  $F_i$ , it can ask for display flatness  $F_i$  according to the following formulas.

$$[0046] F_i = N / \{ |LLi1 - LLi0| + |LLi2 - LLi1|, + |LLi(N-1) - LLi(N-2)| \}$$

[Effect of the Invention] Even if it decodes the coded data of each tile by this invention, the difference of the image quality for every tile is mitigable with the above explanation.

---

## DESCRIPTION OF DRAWINGS

---

[Brief Description of the Drawings]

[Drawing 1] It is drawing showing the functional configuration of the image coding equipment in the 1st operation gestalt of this invention.

[Drawing 2] It is drawing showing the example of processing of tile division.

[Drawing 3] It is drawing in which carrying out tile division of the input image, and showing the functional configuration of the divided conventional image coding equipment which encodes for every tile.

[Drawing 4] It is drawing showing a subband.

[Drawing 5] It is drawing showing the basic configuration of the image coding equipment in the operation gestalt of this invention.

---

(51) Int.Cl.<sup>7</sup>

識別記号

F I

テーマコード\*(参考)

H 0 4 N 7/30  
1/41H 0 4 N 1/41  
7/133B 5 C 0 5 9  
Z 5 C 0 7 8

審査請求 未請求 請求項の数20 O L (全 6 頁)

(21) 出願番号 特願2001-218665(P2001-218665)

(22) 出願日 平成13年7月18日 (2001.7.18)

(71) 出願人 000001007

キヤノン株式会社

東京都大田区下丸子3丁目30番2号

(72) 発明者 大佐 欣也

東京都大田区下丸子3丁目30番2号 キヤ  
ノン株式会社内

(74) 代理人 100076428

弁理士 大塚 康徳 (外3名)

Fターム(参考) 5C059 KK01 MA00 MA23 MA24 MC11

PP01 SS15 SS20 TD03 TD04

UA02 UA15

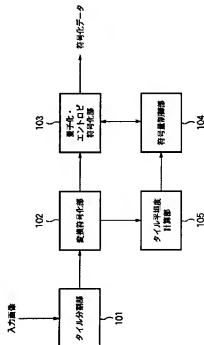
5C078 BA21 BA53 CA21 DA00 DA01

(54) 【発明の名称】 画像符号化装置、及び画像符号化方法、並びにプログラムコード、記憶媒体

(57) 【要約】

【課題】 各タイルの符号化データを復号しても、各タイル毎の画質の差を軽減すること。

【解決手段】 符号量制御部104は、タイル分割部101によって分割されたタイル毎に割り当てる基本的な符号量を決定し、タイル平坦度計算部107によって計算されたタイル毎の平坦度の情報を用いて、割り当てる基本的な符号量をタイル毎に増減する制御を行う。



## 【特許請求の範囲】

【請求項 1】 JPEG2000 (ISO/IEC 15444-1) 規格に従って画像を符号化する画像符号化装置であって、画像をタイルに分割する分割手段と、

各タイルに対して変換処理を施し、変換係数を生成する変換手段と、

前記変換係数に対して符号化処理を行う符号化手段と、各タイル毎に割り当てる基本符号量を計算する基本符号量計算手段と、

注目タイルのデータを元に注目タイルの平坦度を示す指標を計算する指標計算手段と、

前記基本符号量計算手段により計算された注目タイルに割り当てる基本符号量を、前記指標計算手段により計算された注目タイルの平坦度を示す指標に応じて、平坦度が大きいほど割当量が増えるように、割当量を増減させる修正符号量計算手段と、

前記修正符号量計算手段により計算された注目タイルの割り当て符号量に基づいて前記符号化手段を制御する制御手段とを備えることを特徴とする画像符号化装置。

【請求項 2】 前記指標計算手段は、前記注目タイルのデータの分散の逆数を前記注目タイルの指標として用いることを特徴とする画像符号化装置。

【請求項 3】 前記指標計算手段は、前記注目タイルの各データと当該各データの平均値との差分の絶対和を求め、その逆数を前記注目タイルの指標として用いることを特徴とする請求項 1 に記載の画像符号化装置。

【請求項 4】 前記指標計算手段は、前記注目タイルの各データを順にスキャンしたときに、隣合う 2 つのデータの差の絶対値の総和を求め、その逆数を前記注目タイルの指標として用いることを特徴とする請求項 1 に記載の画像符号化装置。

【請求項 5】 前記注目タイルのデータは、前記注目タイルに含まれる画素の値であることを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の画像符号化装置。

【請求項 6】 前記注目タイルのデータは、前記注目タイルの LL サブバンドを構成する変換係数であることを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の画像符号化装置。

【請求項 7】 前記基本符号量計算手段は、前記注目タイルに対して割り当てる基本となる符号量を、レート／ディストーション最適化を用いて求めることを特徴とする請求項 1 に記載の画像符号化装置。

【請求項 8】 前記注目タイルの指標を  $F$ 、前記注目タイルに対して割り当てる基本的な符号量を  $B_{base}$ 、関数  $f(x)$  を  $x$  の値に対して単調に増加する関数とした場合、

前記修正符号量計算手段は、前記注目タイルに対して割り当てる符号量  $B = B_{base} \times (1 + f(F))$  を計算することであることを特徴とする請求項 1 に記載の画像符号化装置。

【請求項 9】 前記修正符号量手段は、 $x$  の大きな値に対して  $f(x)$  が上限値に飽和し、 $x$  の小さな値に対して  $f(x)$  が下限値に飽和することを特徴とする請求項 8 に記載の画像符号化装置。

【請求項 10】 JPEG2000 (ISO/IEC 15444-1) 規格に従って画像を符号化する画像符号化方法であって、画像をタイルに分割する分割工程と、

各タイルに対して変換処理を施し、変換係数を生成する変換工程と、

前記変換係数に対して符号化処理を行う符号化工程と、各タイル毎に割り当てる基本符号量を計算する基本符号量計算工程と、

注目タイルのデータを元に注目タイルの平坦度を示す指標を計算する指標計算工程と、

前記基本符号量計算工程で計算された注目タイルに割り当てる基本符号量を、前記指標計算工程で計算された注目タイルの平坦度を示す指標に応じて、平坦度が大きいほど割当量が増えるように、割当量を増減させる修正符号量計算工程と、

前記修正符号量計算工程で計算された注目タイルの割り当て符号量に基づいて前記符号化工程を制御する制御工程とを備えることを特徴とする画像符号化方法。

【請求項 11】 前記指標計算工程では、前記注目タイルのデータの分散の逆数を前記注目タイルの指標として用いることを特徴とする請求項 10 に記載の画像符号化方法。

【請求項 12】 前記指標計算工程では、前記注目タイルの各データと当該各データの平均値との差分の絶対和を求め、その逆数を前記注目タイルの指標として用いることを特徴とする請求項 10 に記載の画像符号化方法。

【請求項 13】 前記指標計算工程では、前記注目タイルの各データを順にスキャンしたときに、隣合う 2 つのデータの差の絶対値の総和を求め、その逆数を前記注目タイルの指標として用いることを特徴とする請求項 10 に記載の画像符号化方法。

【請求項 14】 前記注目タイルのデータは、前記注目タイルに含まれる画素の値であることを特徴とする請求項 10 乃至 13 のいずれか 1 項に記載の画像符号化方法。

【請求項 15】 前記注目タイルのデータは、前記注目タイルの LL サブバンドを構成する変換係数であることを特徴とする請求項 10 乃至 13 のいずれか 1 項に記載の画像符号化方法。

【請求項 16】 前記符号量計算工程では、前記注目タイルに対して割り当てる基本となる符号量を、レート／ディストーション最適化を用いて求めることを特徴とする請求項 10 に記載の画像符号化方法。

【請求項 17】 前記注目タイルの指標を  $F$ 、前記注目タイルに対して割り当てる基本的な符号量を  $B_{base}$ 、関数  $f(x)$  を  $x$  の値に対して単調に増加

する間数とした場合、

前記修正符号量計算工程では、前記注目タイルに対して割り当てる符号量  $B = B_{base} \times (1 + f(F))$  を計算することで求めることを特徴とする請求項 10 に記載の画像符号化方法。

【請求項 18】 前記修正符号量手段は、 $x$  の大きな値に対して  $f(x)$  が上限値に飽和し、 $x$  の小さな値に対して  $f(x)$  が下限値に飽和することを特徴とする請求項 17 に記載の画像符号化方法。

【請求項 19】 請求項 10 乃至 18 のいずれか 1 項に記載の画像符号化方法を実行するプログラムコード。

【請求項 20】 請求項 19 に記載のプログラムコードを格納し、コンピュータが読み取り可能な記憶媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、画像を符号化する画像符号化装置、及び画像符号化方法、並びにプログラムコード、記憶媒体に関するものである。

【0002】

【従来の技術】JPEG2000 画像符号化方式 (ISO/IEC FDIS15444-1) のように、画像を Wavelet 変換して符号化する符号化方式は、ブロックベースの DCT を利用する JPEG ベースライン符号化方式等と比較して、圧縮度を上げてブロックむずみを生じず画質的なメリットがある。このような画質的なメリットを最大限に生かすには、入力画像全体に対して Wavelet 変換を施すことが有効であるが、実際に符号化装置を実装するにあたっては、入力画像のサイズが非常に大きい、あるいは、符号化装置のバッファ容量が限られている、等の理由により、入力画像全体を一度に Wavelet 変換、符号化せず、入力画像を複数の領域 (タイル) に分割して、各タイルをそれぞれ符号化するという手法を取ることがある。

【0003】このようなタイル分割の処理例を図 2 に示す。同図において入力画像 21 はタイル分割部 24 によって点線部分でタイル 22 やタイル 23 に分割され、分割されたタイル 22、23 は一つずつ符号化器で符号化される。分割されたタイル 22、23 は入力画像 21 と比べてサイズが小さいため、この後の符号化器で持つバッファ容量を限定し、より容易に低コストでの実装を可能にする。

【0004】図 3 に入力画像をタイル分割し、分割されたタイル毎に符号化を行う画像符号化装置の機能構成を示す。

【0005】301 はタイル分割部で、入力画像を所定のサイズを有する領域 (タイル) に分割し、分割したタイルを一つずつ変換符号化部 302 に出力する。302 は変換符号化部で、入力されたタイルに対して変換処理を行い、変換係数を生成し、量子化・エントロピ符号化部 303 に出力する。この変換処理は例えば周波数変換

があり、周波数変換として離散ウェーブレット変換や DCT 等がある。303 は量子化・エントロピ符号化部 303 で、まず変換係数に対して量子化を行い、変換係数の量子化値を得る。そしてこの量子化値に対してエントロピ符号化を行い、符号化データとして出力する。なお、この符号化データにはヘッダが添付され、このヘッダには上述の符号化処理に用いたパラメータ、例えば量子化ステップやタイルのサイズなどが含まれている。304 は符号量制御部で、量子化・エントロピ符号化部 303 において量子化、エントロピ符号化を行う際に発生する符号量を監視しており、発生する符号量に応じて量子化・エントロピ符号化部 303 を制御 (符号化制御) する。例えば、発生する符号化データのサイズが予め指定したサイズになるようにする。

【0006】この符号量制御の手法として、ISO/IEC FDIS15444-1 の "Rate control" の節では、指定された符号量において、符号化された画像と入力画像の誤差を最小化するレート/ディストーション最適化の手法を用いており、誤差が最小という意味で最も良い画質が得られるよう、発生する符号量を制御している。

【0007】

【発明が解決しようとする課題】レート/ディストーション最適化の手法を用いて符号量制御を行うことは、誤差が最小という意味では最良の画質が得られるような制御となる。しかし、人間の視覚的な特性により、誤差が最小であることと視覚的に画質が良いことは必ずしも一致しない。特に、タイル分割を行って符号化を行う場合には、誤差を最小化するだけでは様々な問題を生じる。この問題を図 2 に示した入力画像 21 を用いて説明する。

【0008】まず入力画像 21 において平坦な画像領域を多く含むタイル 22、平坦でない画像領域を多く含むタイル 23 の比較を行う。ここで平坦な画像領域とは、夫々の画素値の変化が比較的小さい画素を含む領域を示す。一方、平坦でない画像領域とは、夫々の画素値の変化は比較的大きい画素を含む領域を示す。

【0009】ISO/IEC FDIS15444-1 に記述された符号量制御を行うと、画像全体に対する誤差を最小化するように符号量制御が行われ、タイル 22、タイル 23 にも誤差の面で最適な符号量の配分が行われる。ところが、実際に復号後の画像を目で見てみると、平坦でないタイル 23 の画質は比較的良好なのに対して、平坦なタイル 22 はタイル 23 に比べてややけた画質になり、画像全体としてみると、平坦部と平坦でない部分の画質に差のある違和感のある画像となる。これは、人間の視覚特性におけるマスキング効果の影響であり、平坦でない画像領域では画像情報自体の変動により符号化誤差がマスキングされて多少大きな誤差でも目立たないのに対し、平坦な領域ではマスキングが働かない

ため小さな誤差でも目立つためである。

【0010】本発明は以上の問題に鑑みて成されたものであり、各タイトルの符号化データを復号しても、各タイトル毎の画質の差を軽減することを目的とする。

【0011】

【課題を解決するための手段】本発明の目的を達成するために、例えば本発明の画像符号化装置は以下の構成を備える。

【0012】すなわち、JPEG2000(ISO/IEC 15444-1)規格に従って画像を符号化する画像符号化装置であって、画像をタイトルに分割する分割手段と、

各タイトルに対して変換処理を施し、変換係数を生成する変換手段と、

前記変換係数に対して符号化処理を行う符号化手段と、各タイトル毎に割り当てる基本符号量を計算する基本符号量計算手段と、

注目タイトルのデータを元に注目タイトルの平坦度を示す指標を計算する指標計算手段と、

前記基本符号量計算手段により計算された注目タイトルに割り当てる基本符号量を、前記指標計算手段により計算された注目タイトルの平坦度を示す指標に応じて、平坦度が大きいほど割当量が増えるように、割当量を増減させる修正符号量計算手段と、

前記修正符号量計算手段により計算された注目タイトルの割り当て符号量に基づいて前記符号化手段を制御する制御手段とを備える。

【0013】

【発明の実施の形態】以下添付図面を参照して、本発明を好適な実施形態に従って詳細に説明する。

【0014】【第1の実施形態】図1に本実施形態における画像符号化装置の機能構成を示す。

【0015】101はタイトル分割部で、入力画像を所定のサイズの領域(タイトル)に分割し、各タイトルは変換符号化部102に出力する。

【0016】102は変換符号化部で、各タイトルに対して変換処理を行い、変換係数を生成する。この変換処理として本実施形態では周波数変換を用いる。なお、この周波数変換にはDCTや離散ウェーブレット変換が含まれるがこのいずれでも良い。

【0017】103は量子化・エントロピ符号化部で、変換符号化部102による変換係数に対して所定の量子化ステップを用いて量子化を行い、量子化された変換係数に対してエントロピ符号化を行う。

【0018】104は符号量制御部で、量子化・エントロピ符号化部で行われる量子化、エントロピ符号化部により発生する符号化データの符号量を監視し、この符号量に基づいて量子化・エントロピ符号化部104を制御する。詳細は後述する。

【0019】105はタイトル平坦度計算部で、タイトル毎の平坦度を計算し、計算した平坦度を符号量制御部10

4に出力する。詳細は後述する。

【0020】また、本実施形態における画像符号化装置の基本構成を図5に示す。

【0021】501はCPUで、RAM502やROM503に格納されたプログラムやデータを用いて画像符号化装置全体の制御を行うと共に、図1に示した機能構成を有する本実施形態の画像符号化処理を実行する。

【0022】502はRAMで、外部記憶装置504や記憶媒体ドライブ509からロードされたプログラムやデータ等を一時的に格納するエリアを備えると共に、CPU501が各種の処理を実行する際に用いるワークエリアを備える。

【0023】503はROMで、画像符号化装置全体の制御を行うプログラムやデータなどを格納する。

【0024】504はハードディスクなどの外部記憶装置で、記憶媒体ドライブ509からロードされたプログラムやデータなどを保存することができる。また、ワークエリアのサイズがRAM502の許容を越えた場合に、越えた分のエリアをファイルとして提供することができる。

【0025】505、506はキーボード、マウスで、各種の指示を画像符号化装置に対して入力することができる。

【0026】507は表示装置で、CRTや液晶画面などにより構成されており、各種の画像情報や文字情報などを表示することができる。

【0027】508は画像入力装置で、スキャナやデジタルカメラ等により構成されており、画像をデジタルデータとして画像符号化装置に入力することができる。なお、画像入力装置は画像符号化装置と接続するためのインターフェースを含む。

【0028】509は記憶媒体ドライブで、CD-ROMやDVD-ROM、フロッピー(登録商標)ディスクなどの記憶媒体からプログラムやデータなどを読み取り、RAM502や外部記憶装置504に出力することができる。

【0029】510は上述の各部を繋ぐバスである。

【0030】以上の構成を備える本実施形態の画像符号化装置の、符号量制御部104、タイトル平坦度計算部105における処理について以下、詳細に説明する。

【0031】符号量制御部104は、例えば、ISO/IEC FDIS15444-11に記載されたレート/ディストーション最適化などの手法により、タイトル毎に割り当てる基本的な符号量を決定する。そして、タイトル平坦度計算部17によって計算されたタイトル毎の平坦度の情報を用いて、割り当てる基本的な符号量をタイトル毎に増減する制御を行う。

【0032】タイトル $i$ の平坦度が $F_i$ で与えられる( $F_i$ が大きいほどタイトル $i$ が平坦である)とすると、タイトル $i$ の基本的な符号量 $R_{base-i}$ を次式によ

うにF<sub>i</sub>に応じて増減させることによって、タイルiに実際に割り当てる符号量B<sub>i</sub>を得ることができる。

【0033】

$$B_i = \text{Base} - i \times (1 + f(F_i))$$

ここで関数f(x)は、xの値に対して単調に増加する関数であり、例えば、

$$f(x) = a(x - b) \quad \text{ただし、} a, b \text{ は定数で、} a > 0$$

が考えられる。この例の関数の場合、定数aの値が、平坦度F<sub>i</sub>の変化量に対する符号割当量B<sub>i</sub>の増減の度合いを決め、平坦度F<sub>i</sub>の値が定数bの値より大きい場合は、f(F<sub>i</sub>) > 0、すなわち、B<sub>i</sub>を増加させ、F<sub>i</sub>の値がbの値より小さい場合は、f(F<sub>i</sub>) < 0、すなわち、B<sub>i</sub>を減少させる。なお、定数a、bの値は、視覚的な改善効果が得られ、かつ画像全体としての符号量が意図する符号量になるように定められる。

【0034】関数f(x)は、単調に増加する関数であれば、この例の関数に限らないが、非常に大きなxの値や、非常に小さなxの値に対して、符号割当量B<sub>i</sub>が極端な値に変化することを避けるため、このようなxに対しては、関数f(x)の値を上限または下限で飽和させることが望ましい。

【0035】タイル平坦度計算部105において計算される平坦度F<sub>i</sub>としては、タイルi内の各画素値の分散の逆数を用いるが、これに限定されるものではなく、タイル内の各画素値と画素値の平均値の差分の絶対値和を求め、その逆数を用いても良い。また、あるいは、タイル内の各画素値を順にスキャンしたときに、隣合う2画素値の差の絶対値の総和を求め、その逆数を平坦度として用いることもできる。

【0036】タイルi内の画素値をP<sub>ij</sub> (0 ≤ j < N-1, Nはタイルi内の画素数)、P<sub>ij</sub>の平均値をP' とすると、平坦度F<sub>i</sub>としてタイルi内の各画素値の分散の逆数を用いた場合、以下の式に従って平坦度F<sub>i</sub>を求めることができる。

$$F_i = N / \{ (P_{i0} - P')^2 + (P_{i1} - P')^2, \dots, + (P_{i(N-1)} - P')^2 \}$$

一方、平坦度F<sub>i</sub>として、タイル内の各画素値と画素値の平均値の差分の絶対値和を求め、その逆数を用いた場合、平坦度F<sub>i</sub>は以下の式に従って求めることができる。

$$F_i = N / \{ |P_{i0} - P'| + |P_{i1} - P'|, \dots, + |P_{i(N-1)} - P'| \}$$

あるいは、平坦度F<sub>i</sub>として、タイル内の各画素値を順にスキャンしたときに、隣合う2つの画素値の差の絶対値の総和を求め、その逆数を用いた場合、平坦度F<sub>i</sub>は以下の式に従って求めることができる。

$$F_i = N / \{ |P_{i1} - P_{i0}| + |P_{i2} - P_{i1}|, \dots, + |P_{i(N-1)} - P_{i(N-2)}| \}$$

そして、タイル平坦度計算部105が上述の通り求めた平坦度F<sub>i</sub>を用いて、符号量制御部104は各タイル毎に実際に割り当てる符号量B<sub>i</sub>を上述の式に基づいて求める。

【0040】以上の処理により全てのタイルに対して符号化処理を施すと、各タイルの符号化結果を含む符号化データを生成する。またこの符号化データには各タイル毎の量子化ステップやエンクロピ符号化に用いるパラメータなどを含むヘッダを添付する。

【0041】以上の説明により、本実施形態における画像符号化装置及び画像符号化方法は、各タイル毎の平坦度を考慮した符号化を行うので、これらのタイルを復号しても各タイル毎の画質の差を軽減することができる。

【0042】[第2の実施形態] 変換符号化部101における周波数変換処理としてウェーブレット変換を用いた場合について説明する。変換符号化部101がタイル毎にウェーブレット変換を行うと、各タイルは図4に示すようなサブバンドに分割される。

【0043】図4においてLLサブバンドの係数は、タイルの縮小画像に相当しているため、タイル全体に対して平坦度の計算をしなくても、LLサブバンドに対して平坦度の計算を行えば、少ない計算量でタイルの平坦度を算出できる。具体的には、タイルiのLLサブバンドの係数値をLL<sub>ijk</sub> (0 ≤ k < M-1, MはLLサブバンド内の変換係数の数)、LL<sub>ijk</sub>の平均値をLL' とすると、平坦度F<sub>i</sub>としてタイルi内の各変換係数の分散の逆数を用いた場合、以下の式に従って平坦度F<sub>i</sub>を求めることができる。

$$F_i = M / \{ (LL_{i0} - LL')^2 + (LL_{i1} - LL')^2, \dots, + (LL_{i(M-1)} - LL')^2 \}$$

一方、平坦度F<sub>i</sub>として、サブバンドLL内の各変換係数と変換係数の平均値の差分の絶対値を求め、その逆数を用いた場合、平坦度F<sub>i</sub>は以下の式に従って求めることができる。

$$F_i = M / \{ |LL_{i0} - LL'| + |LL_{i1} - LL'|, \dots, + |LL_{i(N-1)} - LL'| \}$$

あるいは、平坦度F<sub>i</sub>として、サブバンドLL内の各変換係数値を順にスキャンしたときに、隣合う2つの変換係数値の差の絶対値の総和を求め、その逆数を用いた場合、平坦度F<sub>i</sub>は以下の式に従って求めることができる。

$$F_i = N / \{ |LL_{i1} - LL_{i0}| + |LL_{i2} - LL_{i1}|, \dots, + |LL_{i(N-1)} - LL_{i(N-2)}| \}$$

【発明の効果】以上の説明により、本発明によって、各タイルの符号化データを復号しても、各タイル毎の画質の差を軽減することができる。

【図面の簡単な説明】

【図1】本発明の第1の実施形態における画像符号化装

置の機能構成を示す図である。

【図2】 タイル分割の処理例を示す図である。

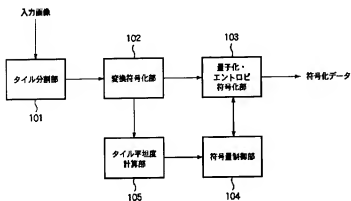
【図3】 入力画像をタイル分割し、分割されたタイル毎に符号化を行う従来装置の機能構成を示す

図である。

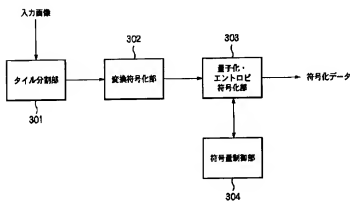
【図4】 サブバンドを示す図である。

【図5】 本発明の実施形態における画像符号化装置の基本構成を示す図である。

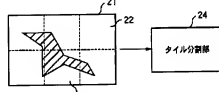
【図1】



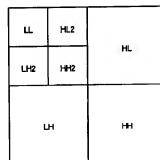
【図3】



【図2】



【図4】



【図5】

